

Requested Patent: JP9069031A

Title:

DISK ARRAY SUBSYSTEM WITH SELF-REALLOCATION OF
LOGICAL VOLUMES FOR REDUCTION OF I/O PROCESSING
LOADS ;

Abstracted Patent: EP0762267, A3 ;

Publication Date: 1997-03-12 ;

Inventor(s):

TAKAMATSU HISASHI (JP); TABUCHI HIDEO (JP); SHIMADA
AKINOBU (JP) ;

Applicant(s): HITACHI LTD (JP) ;

Application Number: EP19960113205 19960816 ;

Priority Number(s): JP19950223029 19950831 ;

IPC Classification: G06F3/06 ;

Equivalents: US5905995 ;

ABSTRACT:

A disk array subsystem includes a logical volume allocation controller (3, 4, 6), which divides and reallocates data of the same logical volume LVN among a plurality of magnetic disk drive groups (12) when a magnetic disk group is added thereto. Successful logical-volume allocation can reduce the I/O processing frequency between magnetic disk drive groups, thus achieving enhanced performance of the disk array subsystem (13).

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-69031

(43) 公開日 平成9年(1997)3月11日

(51) Int. Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 3/06	5 4 0		G 0 6 F 3/06	5 4 0
12/00	5 0 1		12/00	5 0 1 A
12/08		7623-5B	12/08	W

審査請求 未請求 請求項の数3 O L (全 6 頁)

(21) 出願番号 特願平7-223029

(22) 出願日 平成7年(1995)8月31日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 田淵 英夫

神奈川県小田原市国府津2880番地株式会社

日立製作所ストレージシステム事業部内

(72) 発明者 高松 久司

神奈川県小田原市国府津2880番地株式会社

日立製作所ストレージシステム事業部内

(72) 発明者 島田 朗伸

神奈川県小田原市国府津2880番地株式会社

日立製作所ストレージシステム事業部内

(74) 代理人 弁理士 高橋 明夫 (外1名)

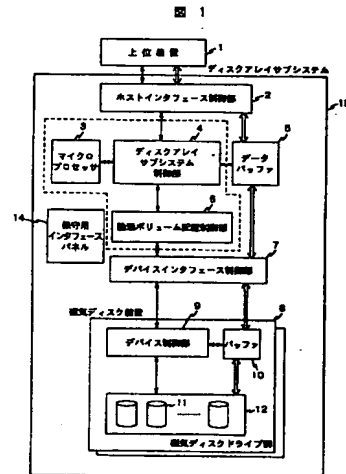
(54) 【発明の名称】 ディスクアレイサブシステム

(57) 【要約】

【目的】 容量拡大に伴う磁気ディスク群の増設時に同一の磁気ディスクドライブ群に対する入出力処理の集中を低減させること。

【構成】 ディスクアレイサブシステム13に、論理ボリューム配置制御部6を設け、磁気ディスク群の増設時に、同一の論理ボリュームLVNのデータを、複数の磁気ディスクドライブ群に分割、再配置する。

【効果】 効率的な論理ボリューム配置により、各磁気ディスクドライブ群間の入出力処理頻度差を軽減でき、ディスクアレイサブシステムの性能の向上が達成できる。



【特許請求の範囲】

【請求項1】複数の磁気ディスクドライブによって構成される磁気ディスクドライブ群と、前記磁気ディスクドライブ群と上位装置とのデータ転送を制御するディスク制御装置とを有するディスクアレイサブシステムにおいて、既に設定されている論理的なボリューム配置を、磁気ディスクドライブの設置変更に応じた論理ボリューム情報に基づき、前記ディスクアレイサブシステム側で自律的に変更し、データの移動を行うことを特徴とするディスクアレイサブシステム。

【請求項2】前記ボリューム配置の変更及びデータの移動は、特定の磁気ディスクドライブへの入出力処理の集中及び入出力処理の競合の発生を防止するように論理的なボリュームの配置が行われることを特徴とする請求項1記載のディスクアレイサブシステム。

【請求項3】前記ボリューム配置の変更及びデータの移動時は、上位装置とディスクアレイサブシステムとの間の処理を継続した状態のまま、ボリュームの配置が行われることを特徴とする請求項1又は請求項2記載のディスクアレイサブシステム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、コンピュータシステムに使用される記憶装置の一つであるディスクアレイサブシステムに係り、特に論理的なボリュームの再配置を自律的に行うディスクアレイサブシステムに関する。

【0002】

【従来の技術】従来より、磁気ディスクサブシステムはコンピュータシステムの外部記憶装置の中心的存在であり、重要な役割を担っている。その中でもディスクアレイサブシステムは、高信頼化を実現した磁気ディスクサブシステムとして注目されてきた。

【0003】ディスクアレイサブシステムはその特長から5つのレベルに分類されている。現在は、5レベルのうち、比較的低コストで高信頼・高可用性が実現でき、小容量ランダムアクセス主体のトランザクション処理に向くとされるRAID5技術が特に注目されている。ディスクアレイサブシステムは、データを記録再生する磁気ディスク駆動装置と、その磁気ディスク駆動装置と上位装置との間のデータ転送を制御する磁気ディスク制御装置等により構成される。

【0004】磁気ディスク駆動装置は、データを記録する磁気ディスクドライブを有しており、RAID5技術では複数の磁気ディスクドライブから成る磁気ディスクドライブ群を一つのグループとし、これを一つ以上の論理的なボリューム（以下、論理ボリューム）に見立てる形態をとる。

【0005】このようなディスクアレイサブシステムにおいてサブシステム容量を拡大する場合には、磁気ディスクドライブを増設し、論理ボリュームを増設、或いは

論理ボリュームの容量を拡大という形態をとる。ところが特定の論理ボリュームに対し、入出力処理が集中する様な環境では、当該論理ボリュームが配置された磁気ディスクドライブ群に入出力処理が集中し、同処理の競合による磁気ディスクドライブの処理待ち時間が増大する。これによってディスクアレイサブシステムの性能を十分に発揮できない恐れがある。

【0006】特開平06-161837号公報には、システムがボリューム選択方式群よりボリュームの選択方式を選択し、自動的に使い分けことが記載されている。ボリュームの選択方式を変更することができるため、特定のボリュームへの入出力負荷の集中を低減するための効率的なファイル配置が実現できる。

【0007】

【発明が解決しようとする課題】情報化社会の進展に伴い、保有するデータ量も増大の一途を辿っている。そのため、磁気ディスク装置にもさらなる大容量化、高信頼化が求められている。そのため、高信頼を実現したディスクアレイサブシステム等が用いられるが、容量拡大のニーズに対しては、従来技術では、磁気ディスクドライブを増設し、その分、論理ボリュームを増設、あるいは論理ボリュームの容量拡大という方法を用いていた。

【0008】また、容量拡大に伴い、特定のファイルに対する入出力処理の集中が懸念されるが、これに関しては従来技術ではシステムでボリュームの選択を自動的に行うことで、効率的なファイルの配置を実現している。

【0009】しかし、ディスクアレイサブシステムにおいて、論理ボリュームの増設等が行われた場合等は、特定の論理ボリュームに対し処理が集中する環境では、従来技術のようなファイル配置方法だけでは、同一磁気ディスクドライブ群への入出力処理の集中を防ぐことは難しく、結果として入出力処理の競合による処理待ち時間の増加により、本来のサブシステム性能が十分に発揮できないことが考えられる。

【0010】そこで本発明の目的は、容量拡大に伴う磁気ディスクドライブ群の増設時に、同一の磁気ディスクドライブ群に対する入出力処理の集中を低減し、磁気ディスクアレイサブシステムのトータル性能の向上を図ることである。

【0011】

【課題を解決するための手段】上記目的を達成するために本発明によれば、ディスクアレイサブシステムにおいて、磁気ディスクドライブ群の増設に伴い、論理ボリュームの増設が行われると、ディスクアレイサブシステムのディスク制御装置は、ファイル属性や全ボリューム数、論理ボリュームと磁気ディスクドライブとの対応関係等の論理ボリューム情報を受け、この論理ボリューム情報に基づき、上位装置との連携なく、自律的に、同一の磁気ディスクドライブ群に対する入出力処理の集中を低減させるように、例えば、各磁気ディスクドライブの

記憶容量域をサブシステム内の全論理ボリューム数に分割し、その分割した各磁気ディスクドライブの各領域に番号を付け、各領域の番号と論理ボリュームの番号が一定条件のもとに合致するように、論理ボリュームの配置変更及びデータの移動を行うようにしたものである。なお、これらの処理は通常処理と並行して行う。

【0012】

【作用】本発明によれば、処理が集中する論理ボリュームが配置された磁気ディスクドライブ群への入出力処理の集中による競合を低減させ、各磁気ディスクドライブ群間における入出力処理頻度の差を小さくすることができる。

【0013】これにより、磁気ディスクドライブの処理待ち時間を低減し、ディスクアレイサブシステムの高性能化を実現することができる。

【0014】また、論理ボリュームの再配置は、上位装置とディスクアレイサブシステムとの通常処理と並行して行うため、通常処理は中止することなく継続できる。

【0015】

【実施例】以下、本発明の実施例について、図面を用いて説明する。第1図は本発明を適用したディスクアレイサブシステムのシステム構成の一例を示すブロック図であり、第2図は本実施例のディスクアレイサブシステムの作用を説く説明図である。

【0016】まず、第1図を参照しながら本発明を適用したディスクアレイサブシステムの構成を説明する。第1図中の単線矢印は制御信号路を示し、複線矢印はデータ路を示す。このディスクアレイサブシステム13は、上位装置1と接続された磁気ディスク制御部と、その配下に接続された磁気ディスク装置8を有し、両者によってRAID5技術を実現したディスクアレイサブシステム13を構成している。磁気ディスク制御部は上位装置1との情報の転送を行うホストインタフェース制御部2、データを一時保管するデータバッファ5、サブシステム13全体の制御を行うマイクロプロセッサ3、マイクロプロセッサ3の指示に基づきコマンドの発行や受領等を行うディスクアレイサブシステム制御部4、論理ボリューム配置の制御を行う論理ボリューム配置制御部6、及び、磁気ディスク装置8との間で情報の転送を行うデバイスインタフェース制御部7を有し、磁気ディスクドライブ11と上位装置1との間のデータ転送を制御する。マイクロプロセッサ3が各種の制御を行う際に使用する制御情報を格納する制御メモリは、本発明とは直接関係しないため省略したが、ここでは、マイクロプロセッサ3に含まれる。また、LSIの性能、規模等に応じて、実際には、図1に破線で囲って示したように、マイクロプロセッサ3、ディスクアレイサブシステム制御部4及び論理ボリューム配置制御部6を、さらには、ホストインタフェース制御部2、データバッファ5及びデバイスインタフェース制御部7までを含めて、1個のLSI

SIで実現することができる。

【0017】また、ディスクサブシステム13内の保守用インタフェースパネル14はこれを通じてユーザ又は保守員が磁気ディスクドライブ11の交換、増設を行う際に利用するものである。

【0018】磁気ディスク装置8は、磁気ディスクドライブ群12とそれを制御するデバイス制御部9及びデータバッファ10を有する。磁気ディスクドライブ群12は複数の磁気ディスクドライブ11から成り、データは複数のブロックに分割され、順次磁気ディスクドライブ11に格納されている。上位装置1からはこの磁気ディスクドライブ群12が論理的に一つのボリュームとして扱われている。

【0019】サブシステム容量の拡大には、ボリューム数の増加といった手段が考えられるが、これを実現するには、新たに磁気ディスクドライブ群12を増設し、上位装置1にて新ボリューム属性の定義を行うことが必要となる。ただ、新設ボリュームと既設のボリュームでは、属性の差異等から、入出力処理の頻度が異なることが予想される。このため、論理ボリューム配置制御部6は、磁気ディスクドライブ11の増設に伴い、上位装置1より磁気ディスクドライブ群12に既に設定された論理ボリューム配置の変更、及びデータの移動を、上位装置1との連携なくディスクアレイサブシステム13で自律的に行うことができる機能を有する。

【0020】この論理ボリューム配置変更、及びデータ移動の一連の動作は、上位装置1にて新設ボリュームが認識された後にディスクアレイサブシステム13にて自律的に行われる他、保守用インタフェースパネル14を通じて、論理ボリューム配置制御部6へ動作内容の指示をすることによっても可能である。また、逆に、上記論理ボリューム配置変更、及びデータ移動動作の抑止も保守用インタフェースパネル14を通じて行うことが可能である。

【0021】なお、これらの処理は上位装置1とディスクアレイサブシステム13の間の通常処理を中止することなく実行される。

【0022】以下、図1に示すディスクアレイサブシステムの処理動作の一例を、第2図を参照しながら説明する。第2図は、前記ボリューム増設に伴う論理ボリュームの再配置動作処理の一例を示す流れ図である。処理ステップP1では、複数の磁気ディスクドライブから成る磁気ディスクドライブ群を論理的に一つのボリューム15と見なし、論理ボリューム番号LVN1を割り当てている。

【0023】いま、容量拡大のため、論理ボリュームを増設する必要性が生じた場合は、処理ステップP2に示すように、新たに磁気ディスクドライブ群を増設し、上位装置からこの磁気ディスクドライブ群を新たな論理ボリューム番号LVN2のボリューム16と設定する。よっ

て、このままでは、論理ボリューム16の増設後は、各磁気ディスクドライブ群15、16ごとに異なるボリューム番号LVN1、LVN2を持つ論理ボリュームが存在する。

【0024】そこで、論理ボリューム配置制御部6は、処理ステップP3に示すように、ファイル属性や全論理ボリューム数、論理ボリュームと磁気ディスクドライブの対応等といった論理ボリューム情報を受け、この論理ボリューム情報に基づき、処理ステップP4に示すように、上位装置により磁気ディスクドライブ群に対して既に設定された論理ボリュームの配置を、ディスクアレイサブシステム側で変更し、また、データの移動も行う。この処理ステップP4における論理ボリュームの配置の変更及びデータの移動は、上記論理ボリューム情報に基づき、各磁気ディスクドライブ間の入出力処理の偏りを低減させるように行われ、例えば、最も簡潔な例としては、図2に示すように、上位装置より既設定の各論理ボリュームを構成する各磁気ディスクドライブ群の領域にまたがるように行われる。

【0025】すなわち、処理ステップP2で、論理ボリューム番号LVN1であった論理ボリューム15は、例えば、前半部と後半部で分けられ、処理ステップP4に示すように、論理ボリューム番号LVN1(FH)の論理ボリューム17と、論理ボリューム番号LVN1(LH)の論理ボリューム19とに分割され、それぞれ、磁気ディスクドライブ群15、16に配置される。

【0026】同様に、処理ステップP2で、論理ボリューム番号LVN2であった論理ボリューム16は、処理ステップP4に示すように、論理ボリューム番号LVN2(FH)の論理ボリューム18と、論理ボリューム番号LVN2(LH)の論理ボリューム20に分割され、それぞれ、磁気ディスクドライブ群15、16に再配置される。この後、この再配置に伴うデータの移動が自動

的に行われる。この配置変更処理及びデータ移動処理は、上位装置とディスクアレイサブシステムの間で行われる通常の処理を中止することなく行われる。

【0027】

【発明の効果】以上述べたように、本発明によれば、ディスクアレイサブシステムにおいて、論理ボリュームの増設等の設置変更があった場合に、ディスク制御装置側で、上位装置との連携なく、自律的に、設置変更後の論理ボリューム情報に基づき、論理ボリュームの再配置を行い、それに伴うデータの移動を行うようにしたので、例えば、各論理ボリュームが全ての磁気ディスクドライブ群にまたがって設定でき、一つの磁気ディスクドライブ群に対し一つの論理ボリュームが設定されている場合に比べ、同一の磁気ディスクドライブ群に対する入出力処理の発生、及び競合を低減し、各磁気ディスクドライブ群間における入出力処理頻度の差を小さくする等、効率的な論理ボリューム配置が実現でき、ディスクアレイサブシステムの性能を向上できる効果がある。

【図面の簡単な説明】

【図1】本発明の一実施例のシステム構成図である。

【図2】論理ボリューム再配置実行動作を示す流れ図である。

【符号の説明】

8…磁気ディスク装置

11…磁気ディスクドライブ

12…磁気ディスクドライブ群

13…ディスクアレイサブシステム

15…論理ボリューム(ボリューム番号1)

16…論理ボリューム(ボリューム番号2)

17…論理ボリューム(ボリューム番号1)の前半部

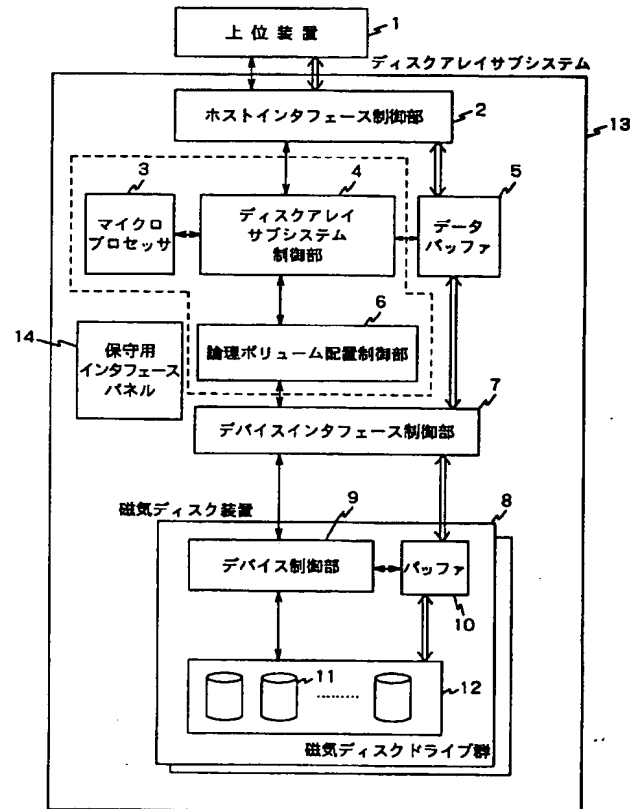
18…論理ボリューム(ボリューム番号2)の前半部

19…論理ボリューム(ボリューム番号1)の後半部

20…論理ボリューム(ボリューム番号2)の後半部

【図1】

図 1



【図2】

図 2

